ADECOR-Net: Improving COVID-19 Lung Infection Segmentation Network with Amplified and Decorrelated Low-level Features

Jiesi Hu, Yanwu Yang, Xutao Guo, Zhikai Chang, Hua Huang, Bo Peng, Ting Ma

Abstract— The Coronavirus Disease 2019 (COVID-19) has become a global pandemic, causing millions of deaths worldwide. Detecting lung infections from computed tomography (CT) images can aid in combating COVID-19. Ground-glass opacities and pneumonic consolidations are the most common features observed in the chest CT scans of infected individuals. These infections are characterized by high heterogeneity and unclear boundaries, making low-level features crucial for segmentation. However, there is a lack of deep learning models designed to explore the low-level features of COVID-19 infection. To address this issue, we propose a network with amplified and decorrelated low-Level features (ADECOR-Net) for precise segmentation of COVID-19 infection. ADECOR-Net comprises a channel re-weighting strategy and a proposed decorrelation loss. The channel re-weighting strategy improves the model's ability to capture richer low-level features without increasing model parameters. The decorrelation loss further reduces the redundancy of captured features and enhances their discriminative power. Comprehensive experiments were conducted, and ADECOR-Net outperforms other methods on two datasets, improving Dice scores by 2.5% and 2.9%. Results highlight its potential for clinical use in precise COVID-19 infection segmentation.

Index Terms— Deep learning, COVID-19 infection segmentation, Low-level feature, Feature decorrelation

I. INTRODUCTION

SINCE its emergence in 2020, the coronavirus disease (COVID-19) has spread to countries worldwide and was

Jiesi Hu , Yanwu Yang, and Xutao Guo are with School of Electronics and Information Engineering, Harbin Institute of Technology at Shenzhen, and The Peng Cheng Laboratory. (e-mail: 405323011@qq.com, yangyanwu1111@gmail.com, 18B952052@stu.hit.edu.cn)
Ting Ma is with School of Electronics and Information Engineering, Harbin Institute of Technology at Shenzhen, The Peng Cheng Laboratory, Guangdong Provincial Key Laboratory of Aerospace Communication and Networking Technology, Harbin Institute of Technology, Shenzhen, and International Research Institute for Artifcial Intelligence, Harbin Institute of Technology, Shenzhen. (e-mail: tma@hit.edu.cn)
Zhikai Chang is with School of Electronics and Information Engineering, Harbin Institute of Technology at Shenzhen. (e-mail: 22B352010@stu.hit.edu.cn)
Hua Huang is with Department of Radiology, The Third People's Hospital of Shenzhen. (e-mail: hncz05hh@163.com)
Bo Peng is with the Institute of Pathogen Biology, Shenzhen Center for Disease Control and Prevention. (e-mail: pb18@163.com)

declared a pandemic by the World Health Organization in March of that year [1], [2]. Despite various vaccination efforts, the number of infected individuals continues to rise as the virus mutates. The pandemic has underscored health disparities within and between countries and will likely have a long-term impact on global society [10]. COVID-19 has presented significant challenges to daily life, the global economy, and public health [4], making testing and diagnostic resources for COVID-19 essential.

Studies have shown that the lungs are the primary target for the virus infection which later spread to other parts of the body [5]. As such, computed tomography (CT) imaging of the lungs is a useful tool for diagnosing and monitoring COVID-19 infection. Several characteristic patterns have been identified in CT scans of COVID-19 patients, including bilateral, peripheral, and basal predominant ground-glass opacity (GGO), multifocal patchy consolidation, and crazy-paving patterns with a peripheral distribution [6]. Ground-glass opacities are defined as an increase in opacification of the lung that does not obscure the blood vessels and airways, while consolidation is defined as a homogeneous opacification that obscures blood vessels and airway walls [7]. The crazy-paving pattern can be explained as GGO with superimposed inter and intralobular septal thickening. Chest CT images can be used to assess the size and severity of these lesions, enabling disease progression monitoring and treatment planning [8]. However, manual annotation of these images is time-consuming and impractical for the numerous suspected and confirmed cases. Thus, the development of reliable AI-aided annotation tools is imperative.

In the field of biomedical image analysis, identifying infected regions is often treated as a segmentation task, with Convolutional Neural Network (CNNs) being the most powerful and commonly used approach today. The superiority of deep learning in computer vision has been well-established over the past few decades [11]. Several deep learning networks were proposed for COVID-19 infection segmentation [12]–[16].

Despite the advances in COVID-19 infection segmentation using deep learning, there are still some challenges that need to be addressed. One challenge is the small inter-class variance between the infected region and the background, as shown in Fig. 1. For instance, GGO boundaries often have low contrast and blurred appearances, making it difficult to identify the infected regions accurately. Thus, subtle textures of the blood
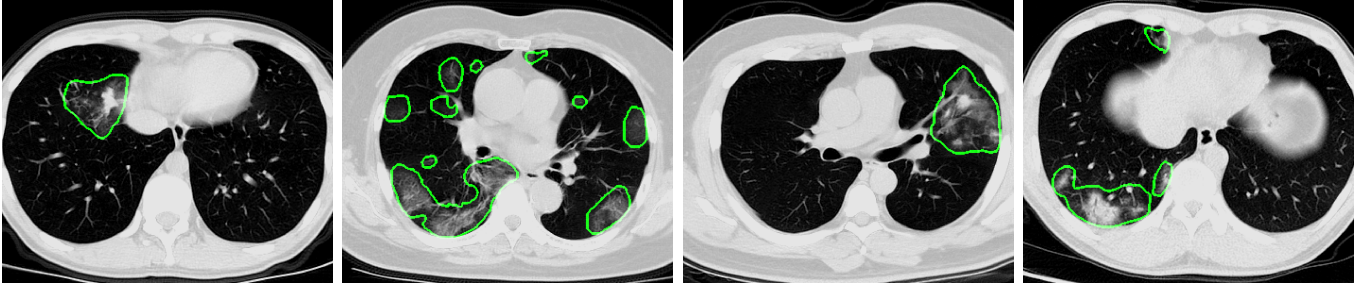
(a) COVID-19 Challenge dataset
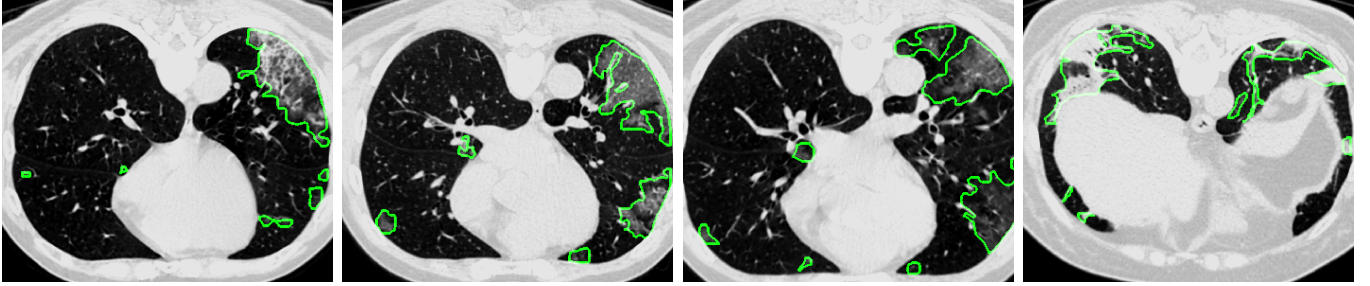
(b) COVID-19 TPHSZ dataset

Fig. 1: (a) and (b) show CT slices from COVID-19 Challenge [17] dataset and the COVID-19 CT dataset collected by the Third People's Hospital of Shenzhen respectivly. The green lines mark the location of the infections.

vessels and airways on the CT slices are the main reference for radiologists when labeling the infected regions [7]. From the perspective of image processing, low-level features such as textures, edges, and intensity are the primary cues for identifying infected regions. Therefore, capturing low-level features is critical for effective segmentation. However, despite the importance of low-level features, most existing models have not fully explored how to better capture them. Additionally, though high-level features can capture a considerable amount of semantic information, they may be less useful for capturing infected regions, which are characterized by high variability in shape and location [15].

In this paper, we propose ADECOR-Net, an improved network with amplified and decorrelating low-Level features, based on our previous work [9]. To relate infected regions precisely, ADECOR-Net increases the model capacity of capturing low-level features by introducing a channel re-weighting strategy that enlarges the shallow layers of the network. Furthermore, to reduce redundancy and increase diversity among captured low-level features, a novel regularization method called decorrelation loss (Decor loss) is introduced. With the help of Decor loss, the extracted low-level features achieve stronger discriminative power, and the model performances are further improved. Experimental results on the COVID-19 Challenge dataset [17] and a private COVID-19 CT dataset demonstrate the effectiveness of our methods.

In this study, we extend our previous work [9] accepted by IEEE International Symposium on Biomedical Imaging 2023 as follow. First, we optimize the normalization layer based on the characteristics of CT slice segmentation (Section III-A). Second, we develop layer-wised decorrelation loss due to various representations from different layers (Section III-C). Third, we validate our methods on an additional dataset collected from diverse scenarios to ensure its robustness and generalizability. Furthermore, to demonstrate the superiority of our model, we conduct comprehensive comparisons with more state-of-the-art models. We also provide extensive experiments and in-depth discussions on the effectiveness of our channel re-weighting strategy and decorrelation loss.

In summary, our paper offers the following contributions:

- We propose a novel strategy for enhancing COVID-19 infection segmentation through channel re-weighting and decorrelating low-level features according to the characteristics of the infection.
- A novel loss function, the Decor loss, is proposed to reduce the correlation among feature channels, via the construction of normalized channel dependencies and the application of cross-entropy in dependency matrices.
- The state-of-the-art performance of our model was substantiated by conducting comprehensive experiments on two datasets, which confirmed the efficacy of enhancing low-level features.

II. RELATED WORKS

In this section, we review the relevant literature on medical image segmentation, COVID-19 infection segmentation, hand-crafted low-level feature, and feature decorrelation methods.

In the field of medical imaging, CNNs have been extensively employed for the analysis of various modalities, including CT, MRI, endoscopic videos, and X-ray images [23]–[26]. Deep learning networks, such as Deeplab [27], U-net [28], fully convolutional network (FCN) [29] and transformer-based models [30], [31], have been proposed and demonstrated promising results. Among these models, U-net has gained significant attention due to its simple and easy-to-train structure. Several modified versions of U-Net, such as Res-UNet [32], UNet++ [33], and Attention U-Net [34], have been developed, and these models have achieved state-of-the-art performance in various medical imaging applications. In this study, we build our model based on Res-UNet, which includes residual connections to avoid the problem of gradient vanishing.

In recent years, several segmentation networks have been proposed for COVID-19 infection segmentation. Inf-Net [12] emphasizes edge information of the lung and introduces specific attention and multi-scale mechanisms. [13] conducted a comprehensive comparison of the performance of U-net and SegNet on COVID-19 two-class and multi-class segmentation tasks. To add spatial attention to the model, [14] proposed a network with multiple attention gates located on the decoder. A contour-aware attention decoder CNN has also been proposed to precisely segment infected tissues by leveraging the crucial boundary and shape information [18]. Based on Unet++, [35] built the squeeze excitation residual module and added atrous spatial pyramid pooling to achieve better performance. CHS-Net [15] developed a hierarchical segmentation model for segregating the coronavirus-infected areas and applied spectral pooling to efficiently reduce the spatial dimension of feature maps. Moreover, [16] proposed a two-stream model that separates the texture and structure components of the image to allow the model to observe different components of the image separately. Although some studies have tried to enhance edge information to achieve better performance, such as Inf-Net [12] and [18], the effectiveness of this approach may be limited due to the blurred boundary of the infections. In our study, we propose an effective way to capture infection characteristics by enhancing the low-level features.

The extraction of low-level features is a fundamental task in computer vision, and various methods have been proposed to achieve this goal. Among them, handcrafted methods have been widely used in many applications. For instance, Laws' filter is a traditional tool for texture classification that relies mainly on low-level features. Its main approach is to filter images with five masks capturing the level, edge, spot, ripple, and wave [36]. The Gabor filters [37] are another effective tool for identifying low-level features [39]. In the field of medical imaging, Gray-Level Co-Occurrence Matrix (GLCM) [41] is also widely used for low-level feature extraction. It is a statistical method of examining texture that considers the spatial relationship of pixels. Despite their effectiveness, integrating these handcrafted features efficiently into deep learning models has been a challenge that remains unsolved.

In this work, we aim to decorrelate the shallow layer of the model to improve the discriminative power of low-level feature maps. Some works in the field of deep learning have been proposed. For instance, Deconv loss [20] and Orth loss [21] are two widely known methods. However, both methods have certain limitations. Deconv loss has the effect of weight decay, which is not always desirable. Orth loss enforces orthogonality in the weight of the convolutional layer, which is not necessarily equivalent to the desired decorrelation of representations. To overcome these challenges, we propose the Decor loss, which achieves better performance in enhancing low-level feature representation.

III. PROPOSED NETWORK

In this section, we present a detailed description of the proposed ADECOR-Net, including its network architecture, core components, and loss function. ADECOR-Net is built on the Res-UNet model, as illustrated in Fig. 2.

The encoder component of the model comprises five Residual Units, and the decoder section includes four Upsample Units. The skip connection enables the decoder to receive spatially detailed information, allowing it to generate fine-grained segmentation. To better capture the infection, the channel re-weighting strategy is proposed and adjusts the number of channels in each layer. Furthermore, we employed the decorrelation loss at the first three layers to reduce the correlation among low-level features, thus minimizing feature redundancy and maximizing diversity of the shallow layers.

A. Residual Unit and Upsample Unit

Our model's Residual Unit comprises two 3×3 convolutional layers. The stride of the first convolutional layer is 2, reducing the input's spatial size by 50%. To retain more spatial information, the fifth Residual Unit's convolutional stride is set to 1, maintaining the spatial size unchanged. The Upsample Unit uses transposed convolution followed by a convolutional layer to extract informative features.

Note that, we replaced the instance norm in the original Res-UNet with batch-norm since the instance norm tends to generate large noise when processing 2D data. In this scenario, using batch-norm instead of instance norm can improve the performance of the baseline model.

B. Channels Re-weighting Strategy

In this section, we describe the channel re-weighting strategy used in our ADECOR-Net to address the limitations of the traditional U-Net and Res-UNet architectures in capturing low-level features. U-Net and Res-UNet typically have five layers with the number of channels increasing as the layers get deeper. However, this design may not be optimal for all scenarios, especially in the case of COVID-19 infection where low-level features play a critical role. These low-level features are essential in identifying COVID-19 infections, as previous studies have shown [6], [42]. Therefore, we modify the channel numbers in the early layers of the network to enhance the model's ability to recognize low-level features. Specifically, we increase the number of channels in the initial layers and gradually decrease them in the deeper layers to keep the model size unchanged and reduce computational requirements.

In our ADECOR-Net, we set the channel numbers from (32, 64, 128, 256, 512) to (248, 248, 112, 112, 112) using
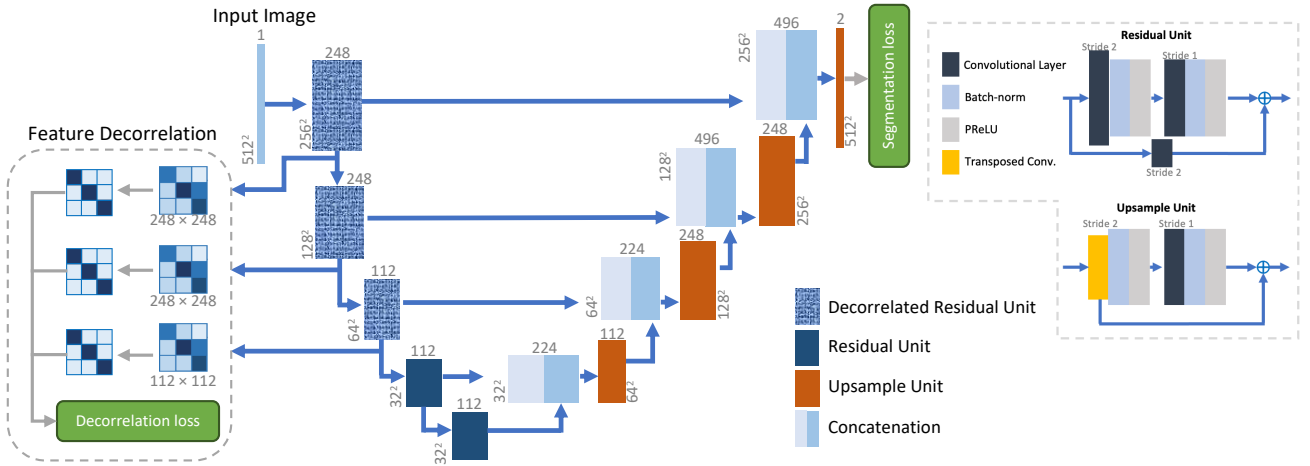
Fig. 2: Overview of the network. The gray numbers around the block are the shape of output feature map.

channel re-weighting. Channel re-weighting will only make ignorable changes to the model parameters and we can directly observe the importance of low-level features to the infection segmentation task. This strategy is verified on both Res-UNet and U-Net architectures. Despite the channel re-weighting, the total number of model layers remains the same as we don't want to shrink the receptive field of the model.
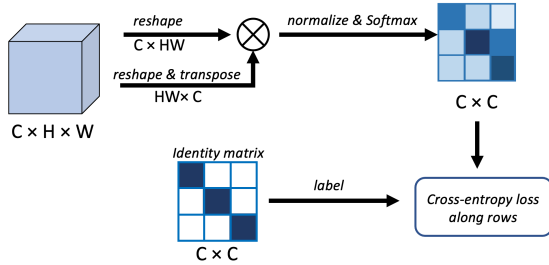

Fig. 3: Pipeline of computing decorrelation loss.

C. Decorrelation Loss

The feature map of a neural network consists of multiple channels, each containing a specific representation of the input. However, increasing the number of channels in shallow layers using the channel re-weighting strategy might result in the redundancy of learned low-level features. To address this issue and ensure the diversity of the low-level features captured by shallow layers, we propose a decorrelation loss.

As illustrated in Figure 3, we calculate the channel correlation map $S \in \mathbb{R}^{C \times C}$ from the original feature map $H \in \mathbb{R}^{C \times H \times W}$ using the channel interdependency approach of [19].

$$s_{i,j} = \sum_{h}^{H} \sum_{w}^{W} h_i^{h,w} h_j^{h,w} \qquad (1)$$

$s_{i,j}$ represents the value in the ith row and jth column of the correlation map S, while $h_i^{h,w}$ represents the value in the hth row and wth column of the ith channel of the feature map

H. To simplify computations, we can reshape $H \in \mathbb{R}^{C \times H \times W}$ to $H \in \mathbb{R}^{C \times HW}$, so that $s_{i,j} = H_i \cdot H_j$. We then apply a softmax function to obtain the probability map $X \in \mathbb{R}^{C \times C}$:

$$x_{i,j} = \frac{exp(\frac{s_{i,j}}{z_i})}{\sum_{k}^{C} exp(\frac{s_{i,k}}{z_i})} \qquad (2)$$

To prevent the model from enlarging the scale of S and reduce the loss, we incorporate the normalization term z_i, which represents the largest value in the ith row of S. We treat the maximum value as a constant when computing the gradient. Afterward, we add the cross-entropy loss to encourage the channel to relate only to itself. The decorrelation loss for a specific feature map is computed as follows:

$$l_{decor} = -\sum_{i}^{C} log(x_{i,i}) \qquad (3)$$

Unlike [20], the proposed Decor loss does not result in weight decay, as increasing or decreasing the scale of the weight does not change X.

Let l_{decor}^{m} denote the Decor loss generated by the feature map of the mth layer. The total Decor loss is as follows:

$$L_{decor} = \sum_{m}^{M} l_{decor}^{m} \qquad (4)$$

where M contains layers with decorrelation loss added. We believe that adding Decor loss to all layers is unnecessary, as semantic information is gradually compressed with the depth of the layer. Therefore, we also investigate the effect of adding Decor loss in different layers. In ADECOR-Net, we only apply the Decor loss to the first three Residual Units, which is sufficient.

D. Loss Function

Our segmentation model utilizes a combination of binary cross-entropy loss (L_{BCE}), dice loss (L_{DC}) and Decor loss (L_{decor}). The task involves classifying each pixel as either infection (positive) or background (negative), and we treat it

as a pixel-wise binary classification problem. The binary cross-entropy loss, as defined in Equation (5), is commonly used in classification and segmentation tasks. It works effectively with balanced sample distribution and can learn the probability of each pixel efficiently.

$$L_{BCE} = -\sum_i^N (y_i log(P(x_i)) + (1 - y_i)log(1 - P(x_i))) \quad (5)$$

where i and N indicate the pixel index and the total number of pixels in an image. y_i and $P(x_i)$ represent the ground truth value and predicted probability of the ith pixel respectively.

However, due to the imbalanced sample distribution between foreground and background pixels, the model requires the Dice loss, as defined in Equation (6), to give equal importance to both classes. The Dice Score Coefficient is also used to evaluate the segmentation performance. The definition of Dice loss is as follows:

$$L_{DC} = 1 - \frac{2\sum_i^N y_i f(x_i) + \epsilon}{\sum_i^N |y_i| + \sum_i^N |f(x_i)| + \epsilon} \quad (6)$$

where $f(x_i)$ presents the predicted result. ϵ is a small constant that prevents numerical instability. The overall loss function L is defined as follows:

$$L = \frac{1}{2}L_{BCE} + \frac{1}{2}L_{DC} + \lambda L_{decor} \quad (7)$$

where $\lambda > 0$ is the weight of the decorrelation loss.

IV. EXPERIMENTS

A. Dataset Description

Our experimental evaluation of the ADECOR-Net algorithm includes two datasets, namely the public COVID-19 Challenge [17] dataset and the COVID-19 CT dataset collected by the Third People's Hospital of Shenzhen (COVID-19 TPHSZ). The COVID-19 TPHSZ dataset consists of 68 CT volumes that have been meticulously annotated slice by slice. The annotated regions, which include GGO and consolidations, are labeled as 1, and the rest of the regions are labeled as 0. The task of segmenting infections is a two-class segmentation problem, and the segmentation networks are trained and evaluated using 2D CT slices extracted from 3D volumes of the CT images. All 2D slices are resized to $512 \times 512 \times 1$. To ensure that the trained models generalize well, both datasets were divided into training, validation, and test sets as shown in Table I. During the dataset splitting process, we ensured that slices from the same volume were kept in the same set to prevent information leakage. In Fig. 1, we present some representative samples from both datasets to showcase their characteristics.

Detail of COVID-19 TPHSZ dataset: This study was approved by the Ethics Committee of The Third People's Hospital of Shenzhen (IRB No. 2022-123). Data were collected from COVID-19 patients who were admitted to the Third People's Hospital of Shenzhen, the only authorized referral hospital for COVID-19 patients in Shenzhen City, between January 11, 2020, and February 21, 2020. All patients underwent a chest CT scan in supine posture when holding breath after a deep breath. The CT scan was taken from the entrance of the chest

cavity to the lower edge of the diaphragm. All CT scans were acquired from Toshiba Aquilion TSX-101A(Toshiba Medical Systems) and Ingenuity FIex (Philips Medical Systems) CT scanners. When labeling, all CT images were independently reviewed by two radiologists (C. D. and D. D.), and another senior radiologist (B.B., >30 years' experience) made the final decision in cases of disagreement.

B. Implementation Details

Our model was implemented using the MONAI framework in PyTorch and was trained on high-performance environments supported by Tesla V100 GPUs. The distribution of training, validation, and test sets is shown in Table I. The model parameters were optimized using the training set, and the validation set was used to select the best model and its probability threshold. The test set was used to evaluate the model's performance. Given a large number of slices without infections, 80% of the slices without annotations in the training set were removed to improve training efficiency. We did not delete any slices from the validation and test sets to ensure the reliability of our results.

We used stochastic gradient descent (SGD) and Adam optimizer [44] during training to optimize the model. The learning rate was initialized to 1×10^{-4} and reduced by a factor of 5 whenever the training loss did not decrease by at least 5×10^{-3} within the last 30 epochs. All models were trained for 300 epochs to ensure convergence. To remove outliers, the intensity of input images was clipped to the [0.5, 99.5] percentiles, followed by z-score normalization. We also applied various augmentation methods, including random rotation, scaling, elastic deformations, gamma correction, mirroring, and intensity shifting. The final value of λ was determined to be $10^{-2.5}$ based on the results of the validation set.

We evaluated our models using four widely adopted metrics: Precision, Recall, Dice coefficient, and Jaccard Index (JA). Note that these metrics were evaluated using 3D volumes reconstructed from 2D slices. These metrics can be computed given the true positive (TP), true negative (TN), false positive (FP), and false negative (FN) predictions as follow:

$$Precision = \frac{TP}{TP + FP} \quad (8)$$

$$Recall = \frac{TP}{TP + FN} \quad (9)$$

$$Dice = \frac{2TP}{2TP + FP + FN} \quad (10)$$

$$JA = \frac{TP}{TP + FP + FN} \quad (11)$$

C. Segmentation Results and Discussion

Quantitative Results: Our study presents a comparison of the proposed segmentation method with other state-of-the-art models, as presented in Table II. All methods were trained using the same protocol, except for Inf-Net, which was trained with its own implementation [12]. We explored the potential

TABLE I: CT slices distribution of two datasets.

Dataset	Total Volumes	Dimension	Training		Validation		Test	
			Volumes	Slices	Volumes	Slices	Volumes	Slices
COVID-19 Challenge [17]	199	$512 \times 512 \times d$	127	3950	32	3172	40	2582
COVID-19 TPHSZ	68	$512 \times 512 \times d$	42	7596	11	4231	15	5562

TABLE II: Performance comparison with state-of-the-art methods on two datasets.

Method	Param.	COVID-19 Challenge dataset				COVID-19 TPHSZ dataset			
		Dice	JA	Precision	Recall	Dice	JA	Precision	Recall
UNet [28]	2.637M	0.6262	0.4860	0.7016	0.6446	0.6076	0.4580	0.6322	0.6173
Inf-Net [12]	33.122 M	0.6129	0.4689	0.6504	0.6508	0.4949	0.3469	0.5342	0.5056
Unet++ [33]	9.045 M	0.5977	0.4580	0.6702	0.6169	0.5778	0.4302	0.6281	0.5846
Unet++ (large) [33]	36.165 M	0.6053	0.4664	0.6672	0.6372	0.5800	0.4354	0.6250	0.6107
Attention U-net [34]	8.725 M	0.5883	0.4500	0.6449	0.6195	0.5772	0.4334	0.5992	0.6312
Attention U-net (large) [34]	34.877 M	0.6296	0.4874	0.6714	0.6706	0.5851	0.4412	0.6118	0.6361
Swin-Unet [30]	41.342 M	0.5998	0.4567	0.6423	0.6502	0.5886	0.4444	0.6408	0.6074
Swin-Unet (pretrained) [30]	41.342 M	0.6356	0.4956	0.6975	0.6650	0.5991	0.4522	0.6299	0.6334
Res-UNet [32]	6.495 M	0.6436	0.5051	0.6815	0.6812	0.6058	0.4565	0.6322	0.6252
Ours	6.457 M	**0.6687**	**0.5297**	**0.7070**	**0.6990**	**0.6349**	**0.4855**	**0.6501**	**0.6599**

of SOTA methods by implementing Unet++ and Attention U-net models of different sizes. In addition, we used a pre-trained model to initialize the parameters of Swin-Unet due to the transformer's high dependency on pre-trained models, which was not done for other methods. The proposed method outperforms other cutting-edge methods by a large margin in all evaluation metrics on both datasets. We attribute this improvement to our channel re-weighting strategy and the proposed Decor loss, which enable our model to effectively explore low-level features. Moreover, our model parameters are of medium size, indicating that it can efficiently make use of model parameters, thereby reducing the risk of overfitting.

Qualitative Results: The lung infection segmentation results are compared in Fig. 4, where the green line represents the ground truth, and the red line represents the prediction of The proposed method. The results in the first, second, and fourth rows show that our model performs better at the edge of the infection. Moreover, in the third row, our method can detect most of the infections, even when they are not very clear, whereas other methods fail to do so. Overall, Fig. 4 demonstrates that our network achieves superior segmentation results, particularly for unclear regions. We attribute this improvement to our method's stronger ability to capture subtle texture information.

Effectiveness of channel re-weighting strategy: We conducted a series of experiments to investigate the impact of the channel re-weighting strategy (CR) on the model's performance by gradually adjusting the number of channels. To better demonstrate the effect of channel re-weighting, we did not apply the Decor loss in this experiment. Fig. 5 shows the change in the Dice coefficient due to channel re-weighting on both datasets. The results indicate that increasing the number of shallow channels significantly improves the model's Dice coefficient (approximately 2%) for both Res-UNet and U-Net with this simple strategy. Notably, different channel settings have similar parameter sizes, suggesting that the improvement of the model does not come from the increase in parameters. However, as the number of low-level features increases and the number of high-level features decreases, the numerical growth of Dice tends to reach a saturation point. The Dice

coefficient of U-Net even decreases when the number of channels in the first layer reaches 240. This finding suggests that more low-level features do not always lead to better performance, and some high-level features should be retained. Although CR significantly improves the model's performance, there is no fixed channel setting that is universally applicable to both models. Therefore, when applying CR to different backbones, the optimal channel setting should be considered a hyperparameter. Overall, this figure highlights the importance of low-level features in infection segmentation.

Effectiveness of the Decor loss: Figure 8 demonstrates the effectiveness of Decor loss in improving model performance under different channel settings. Regardless of the channel setting, the model's performance is consistently enhanced on both datasets when Decor loss is added. However, the degree of improvement varies across different channel settings, with some settings showing an improvement of approximately 1% in Dice. Table III compares the effectiveness of different decorrelation methods. The results indicate that the addition of any decorrelation method improves model performance. Nonetheless, our Decor loss still outperforms the other two methods in terms of Dice and JA on both datasets. Furthermore, we present Figure 7, which illustrates the average channel dependency across the validation set before and after the addition of Decor loss. As can be seen from the figure, the correlation among feature maps is noticeably reduced after the application of Decor loss, which is consistent with the observed improvement in model performance. This finding indicates that the Decor loss not only enhances the performance of the model but also promotes feature diversity by reducing the correlation among feature maps.

D. Sensitivity Analysis

Fig 8 shows the sensitivity of Decor loss to the weight and added layers. Results show that the model performs best when the weight of Decor loss is set to $10^{-2.5}$, whereas using other weights may not guarantee a strong improvement in the model's performance. This is because a strong Decor loss can cause the model to sacrifice fitting accuracy in order to reduce
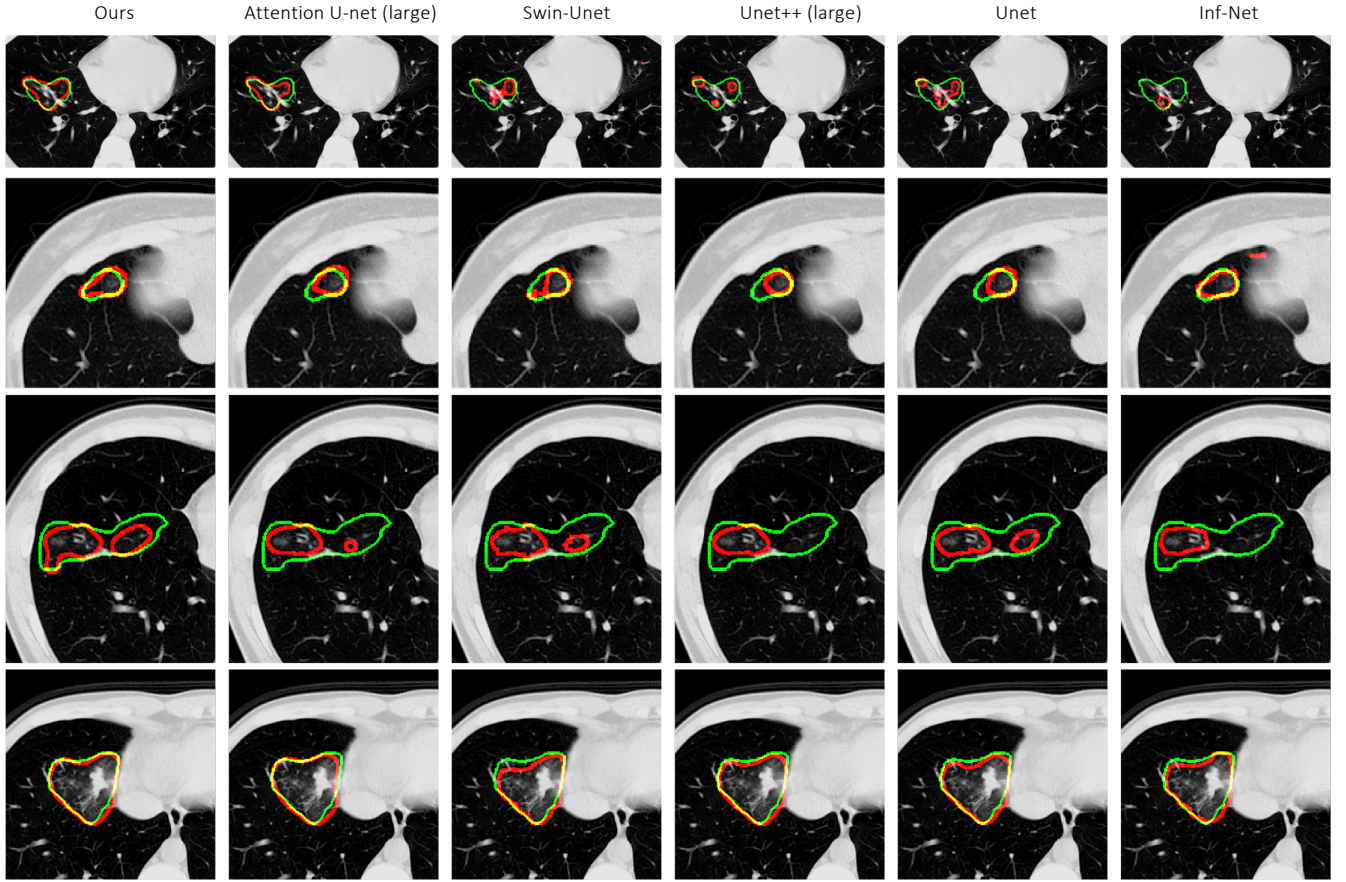
Fig. 4: Visual comparison of the segmentation results. The green line represents the ground truth and the red line represents the prediction made by a specific model.
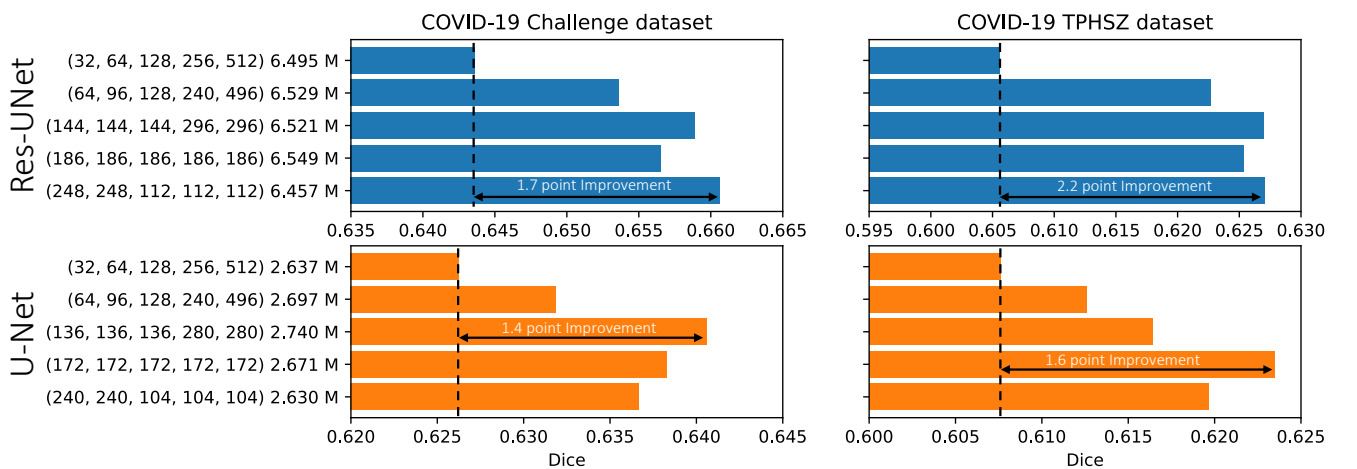

Fig. 5: Impact of channel re-weighting on the Dice coefficient of two datasets. The first row shows the results based on Res-UNet, and the second row shows the results based on the original U-Net. The labels on the y axis denote the channel settings and the number of parameters. The proposed channel re-weighting strategy leads to improved segmentation performance with a slight change in the number of model parameters.
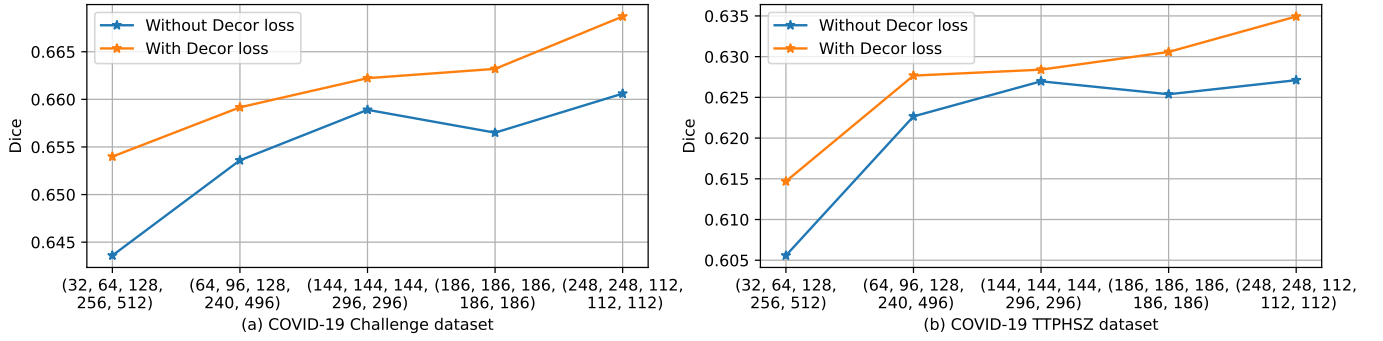
Fig. 6: Effectiveness of Decor loss under different channel settings based on Res-UNet. The labels on the x axis represent the different channel settings.

TABLE III: Model performances using different decorrelation methods, where CR denotes the channel re-weighting strategy.

| Decorrelation Method	COVID-19 Challenge dataset				COVID-19 TPHSZ dataset			
	Dice	JA	Precision	Recall	Dice	JA	Precision	Recall
Only CR	0.6606	0.5216	0.6928	0.6987	0.6271	0.4771	0.6339	0.6636
CR + Deconv loss [20]	0.6641	0.5253	0.7052	0.6945	0.6282	0.4794	**0.6524**	0.6508
CR + Orth loss [21]	0.6639	0.5265	**0.7092**	0.6906	0.6286	0.4811	0.6247	**0.6861**
CR + Decor loss (ours)	**0.6687**	**0.5297**	0.7070	**0.6990**	**0.6349**	**0.4855**	0.6501	0.6599
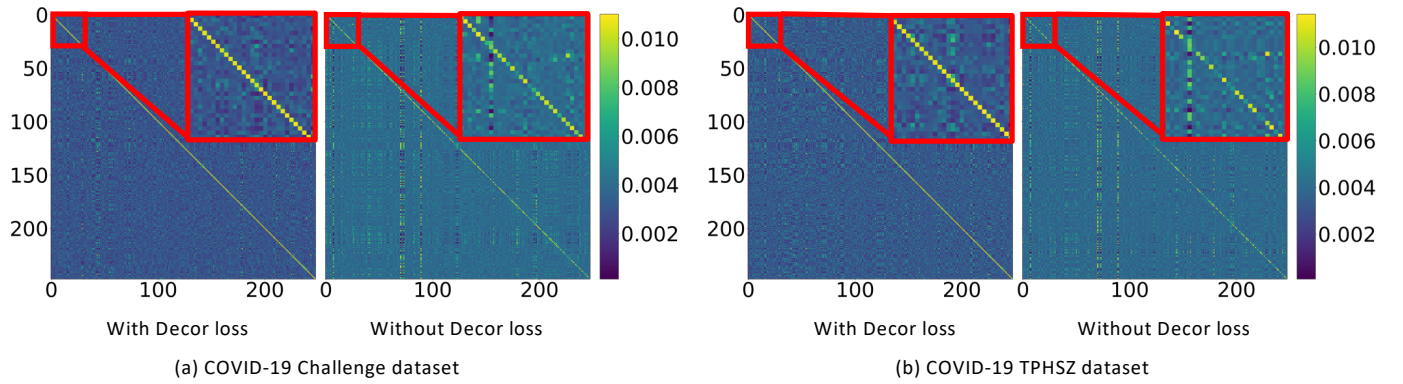
Fig. 7: (a) and (b) show the probability matrix X for two datasets, which illustrates the reduced dependency among different channels using Decor loss.

feature decorrelation, and a very low weight may render the Decor loss ineffective. Additionally, we explored the impact of adding Decor loss to different encoder layers to determine the optimal configuration. Results show that adding Decor loss to the first 3 layers is sufficient to achieve improvement, as Decor loss is designed to obtain diverse low-level features in shallow layers. Therefore, we only added Decor loss to the first 3 layers after applying CR, as it was deemed unnecessary to add Decor loss to deep layers with a relatively small number of channels.

E. Ablation Study

Table IV presents the results of the ablation study for the proposed method on both datasets. The ablation study was conducted under the same training protocol for all models. The results show that both CR and Decor loss can considerably improve the model performance. Specifically, CR increases the Dice coefficient by approximately 2%, and Decor loss results

in a 1% improvement. Moreover, the application of both the Decor loss and CR strategy can further improve the performance on both datasets. This indicates that the improvement produced by CR and Decor loss is complementary and suitable for being applied together.

TABLE IV: Ablation study of proposed method on both datasets.

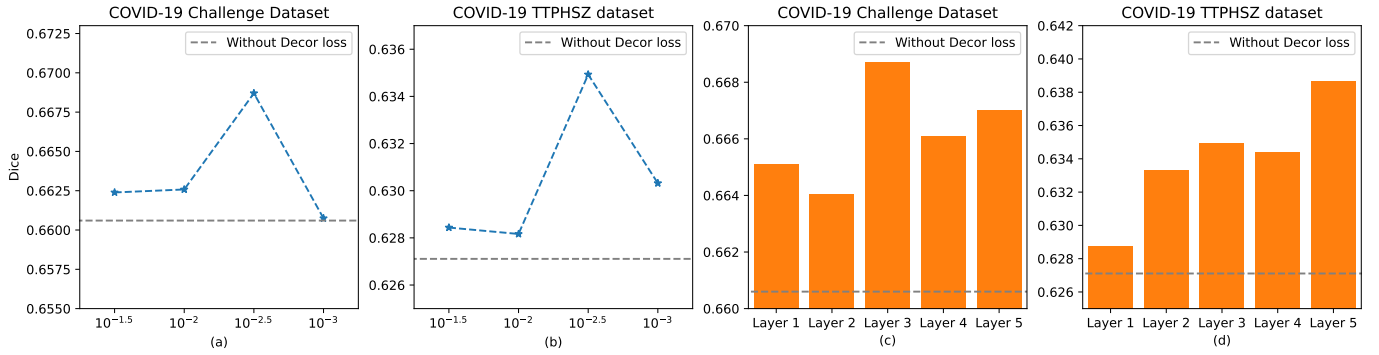| Dataset	CR	Decor loss	Param.	Dice	JA	Precision	Recall
COVID-19 Challenge			6.495 M	0.6436	0.5051	0.6815	0.6812
	✓		6.457 M	0.6606	0.5216	0.6928	0.6987
		✓	6.457 M	0.6540	0.5141	0.6985	0.6856
	✓	✓	6.457 M	**0.6687**	**0.5297**	**0.7070**	**0.6990**
COVID-19 TPHSZ			6.495 M	0.6058	0.4565	0.6322	0.6252
	✓		6.457 M	0.6271	0.4771	0.6339	**0.6636**
		✓	6.457 M	0.6147	0.4664	0.6348	0.6442
	✓	✓	6.457 M	**0.6349**	**0.4855**	**0.6501**	0.6599

Fig. 8: (a) and (b) show the Dice coefficient of the model under different weights of Decor loss. (c) and (d) show the Dice coefficient of the model when Decor loss is added to different encoder layers. For instance, Layer 5 on the x axis label represents adding Decor loss to the first 5 layers of the encoder.

F. Limitations and Future Work

The proposed ADECOR-Net can be further improved by incorporating semi-supervised learning frameworks and using powerful pre-trained models. Semi-supervised learning can enable the model to learn from a large amount of unlabeled data, and a powerful pre-trained model can further improve the feature quality and generalization ability of the model. In addition, we believe that paying more attention to low-level features can also benefit COVID-19 diagnostic models. Therefore, in the future, it would be worthwhile to investigate the performance of the proposed method on the diagnostic task to improve diagnostic accuracy.

V. CONCLUSION

In this paper, our proposed model, ADECOR-Net, demonstrates the importance of emphasizing low-level features of lung images to identify the COVID-19 infection. This is consistent with the way radiologists label the infection. ADECOR-Net involves the channel re-weighting strategy and Decor loss to capture subtle low-level features, which we believe are critical for precise infection segmentation. Our experiments show that the channel re-weighting strategy can effectively capture more low-level features, leading to a significant improvement in performance without increasing the model size. Additionally, we propose the Decor loss to ensure feature diversity while capturing more low-level features, which further improves the overall performance of the model. We compare the proposed ADECOR-Net with other cutting-edge methods and show that it outperforms them in most metrics. Since the proposed methods are easy to implement, they have great potential to be applied in similar scenarios to achieve better performance.

ACKNOWLEDGMENT

REFERENCES

[1] Roosa, Kimberlyn, et al. "Real-time forecasts of the COVID-19 epidemic in China from February 5th to February 24th, 2020." Infectious Disease Modelling 5 (2020): 256-263.

[2] Yan, Li, et al. "Prediction of criticality in patients with severe Covid-19 infection using three clinical features: a machine learning-based prognostic model with clinical data in Wuhan." MedRxiv (2020): 2020-02.

[3] Robinson, Philip C., et al. "COVID-19 therapeutics: Challenges and directions for the future." Proceedings of the National Academy of Sciences 119.15 (2022): e2119893119.

[4] Ucar, Ferhat, and Deniz Korkmaz. "COVIDiagnosis-Net: Deep Bayes-SqueezeNet based diagnosis of the coronavirus disease 2019 (COVID-19) from X-ray images." Medical hypotheses 140 (2020): 109761.

[5] Meredith Wadman, Jennifer Couzin-Frankel, and Catherine Matacic Jocelyn Kaiser. "How does coronavirus kill? Clinicians trace a ferocious rampage through the body, from brain to toes." Science (2020).

[6] Liu, Peng, and Xian-zheng Tan. "2019 novel coronavirus (2019-nCoV) pneumonia." Radiology 295.1 (2020): 19-19.

[7] Ng, Ming-Yen, et al. "Imaging profile of the COVID-19 infection: radiologic findings and literature review." Radiology: Cardiothoracic Imaging 2.1 (2020): e200034.

[8] Shi, Heshui, et al. "Radiological findings from 81 patients with COVID-19 pneumonia in Wuhan, China: a descriptive study." The Lancet infectious diseases 20.4 (2020): 425-434.

[9] Hu, Jiesi, et al. "DECOR-NET: A COVID-19 Lung Infection Segmentation Network Improved by Emphasizing Low-level Features and Decorrelating Features." arXiv preprint arXiv:2302.14277 (2023).

[10] Robinson, Philip C., et al. "COVID-19 therapeutics: Challenges and directions for the future." Proceedings of the National Academy of Sciences 119.15 (2022): e2119893119.

[11] LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. "Deep learning." nature 521.7553 (2015): 436-444.

[12] Fan, Deng-Ping, et al. "Inf-net: Automatic covid-19 lung infection segmentation from ct images." IEEE Transactions on Medical Imaging 39.8 (2020): 2626-2637.

[13] Saood, Adnan, and Iyad Hatem. "COVID-19 lung CT image segmentation using deep learning methods: U-Net versus SegNet." BMC Medical Imaging 21.1 (2021): 1-10.

[14] Budak, Ümit, et al. "Efficient COVID-19 segmentation from CT slices exploiting semantic segmentation with integrated attention mechanism." Journal of Digital Imaging 34 (2021): 263-272.

[15] Punn, Narinder Singh, and Sonali Agarwal. "Chs-net: A deep learning approach for hierarchical segmentation of covid-19 via ct images." Neural Processing Letters 54.5 (2022): 3771-3792.

[16] Elharrouss, Omar, Nandhini Subramanian, and Somaya Al-Maadeed. "An encoder–decoder-based method for segmentation of COVID-19 lung infection in CT images." SN Computer Science 3 (2022): 1-12.

[17] Roth, Holger R., et al. "Rapid artificial intelligence solutions in a pandemic—The COVID-19-20 Lung CT Lesion Segmentation Challenge." Medical image analysis 82 (2022): 102605.

[18] Karthik, R., et al. "Contour-enhanced attention CNN for CT-based COVID-19 segmentation." Pattern Recognition 125 (2022): 108538.

[19] Fu, Jun, et al. "Dual attention network for scene segmentation." Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2019.

[20] Cogswell, Michael, et al. "Reducing overfitting in deep networks by decorrelating representations." arXiv preprint arXiv:1511.06068 (2015).

[21] Wang, Chien-Yao, et al. "CSPNet: A new backbone that can enhance learning capability of CNN." Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops. 2020.

[22] Hu, Jie, Li Shen, and Gang Sun. "Squeeze-and-excitation networks." Proceedings of the IEEE conference on computer vision and pattern recognition. 2018.

[23] Kermany, Daniel S., et al. "Identifying medical diagnoses and treatable diseases by image-based deep learning." cell 172.5 (2018): 1122-1131.

[24] Gómez, Pablo, et al. "Low-light image enhancement of high-speed endoscopic videos using a convolutional neural network." Medical & biological engineering & computing 57 (2019): 1451-1463.

[25] Choe, Jooae, et al. "Deep learning–based image conversion of CT reconstruction kernels improves radiomics reproducibility for pulmonary nodules or masses." Radiology 292.2 (2019): 365-373.

[26] Punn, Narinder Singh, and Sonali Agarwal. "Automated diagnosis of COVID-19 with limited posteroanterior chest X-ray images using fine-tuned deep neural networks." Applied Intelligence 51.5 (2021): 2689-2702.

[27] Chen, Liang-Chieh, et al. "Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs." IEEE transactions on pattern analysis and machine intelligence 40.4 (2017): 834-848.

[28] Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation." Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18. Springer International Publishing, 2015.

[29] Long, Jonathan, Evan Shelhamer, and Trevor Darrell. "Fully convolutional networks for semantic segmentation." Proceedings of the IEEE conference on computer vision and pattern recognition. 2015.

[30] Cao, Hu, et al. "Swin-unet: Unet-like pure transformer for medical image segmentation." Computer Vision–ECCV 2022 Workshops: Tel Aviv, Israel, October 23–27, 2022, Proceedings, Part III. Cham: Springer Nature Switzerland, 2023.

[31] Chen, Jieneng, et al. "Transunet: Transformers make strong encoders for medical image segmentation." arXiv preprint arXiv:2102.04306 (2021).

[32] Diakogiannis, Foivos I., et al. "ResUNet-a: A deep learning framework for semantic segmentation of remotely sensed data." ISPRS Journal of Photogrammetry and Remote Sensing 162 (2020): 94-114.

[33] Zhou, Zongwei, et al. "Unet++: A nested u-net architecture for medical image segmentation." Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 20, 2018, Proceedings 4. Springer International Publishing, 2018.

[34] Oktay, Ozan, et al. "Attention u-net: Learning where to look for the pancreas." arXiv preprint arXiv:1804.03999 (2018).

[35] Xiao, Hanguang, et al. "SAUNet++: an automatic segmentation model of COVID-19 lesion from CT slices." The Visual Computer (2022): 1-14.

[36] Dash, Sonali, and Uma Ranjan Jena. "Multi-resolution Laws' Masks based texture classification." Journal of applied research and technology 15.6 (2017): 571-582.

[37] Gabor, Dennis. "Theory of communication. Part 1: The analysis of information." Journal of the Institution of Electrical Engineers-part III: radio and communication engineering 93.26 (1946): 429-441.

[38] Jain, Anil K., Nalini K. Ratha, and Sridhar Lakshmanan. "Object detection using Gabor filters." Pattern recognition 30.2 (1997): 295-309.

[39] Huang, Lin-Lin, Akinobu Shimizu, and Hidefumi Kobatake. "Classification-based face detection using Gabor filter features." Sixth IEEE International Conference on Automatic Face and Gesture Recognition, 2004. Proceedings.. IEEE, 2004.

[40] Alekseev, Andrey, and Anatoly Bobe. "GaborNet: Gabor filters with learnable parameters in deep convolutional neural network." 2019 International Conference on Engineering and Telecommunication (EnT). IEEE, 2019.

[41] Mohanaiah, P., P. Sathyanarayana, and L. GuruKumar. "Image texture feature extraction using GLCM approach." International journal of scientific and research publications 3.5 (2013): 1-5.

[42] Zeiler, Matthew D., and Rob Fergus. "Visualizing and understanding convolutional networks." Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part I 13. Springer International Publishing, 2014.

[43] He, Kaiming, et al. "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification." Proceedings of the IEEE international conference on computer vision. 2015.

[44] Ruder, Sebastian. "An overview of gradient descent optimization algorithms." arXiv preprint arXiv:1609.04747 (2016).